Journal of Computational and Applied Mathematics 228 (2009) 424–433

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A modified SSOR iterative method for augmented systems[☆]

Shi-Liang Wu, Ting-Zhu Huang*, Xi-Le Zhao

School of Applied Mathematics, University of Electronic Science and Technology of China, Chengdu, Sichuan, 610054, PR China

ARTICLE INFO

Article history:
Received 19 May 2007
Received in revised form 23 March 2008

MSC:
65F10

Keywords:
SOR-like method
SSOR method
Augmented system

ABSTRACT

Golub, Wu and Yuan [G.H. Golub, X. Wu, J.Y. Yuan, SOR-like methods for augmented systems, BIT 41 (2001) 71–85] have presented the SOR-like algorithm to solve augmented systems. In this paper, we present the modified symmetric successive overrelaxation (MSSOR) method for solving augmented systems, which is based on Darvishi and Hessari's work above. We derive its convergence under suitable restrictions on the iteration parameter, determine its optimal iteration parameter and the corresponding optimal convergence factor under certain conditions. Finally, we apply the MSSOR method to solve augmented systems.

1. Introduction

In this paper, the iterative method to solve augmented systems

$$\begin{pmatrix} A & B \\ B^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ q \end{pmatrix} \tag{1}$$

is considered, where $A \in R^{m\times m}$ is symmetric and positive definite, $B \in R^{m\times n}$, $n \le m$ is of column full rank. Under these assumptions, (1) has a unique solution. Systems such as (1) are important and appear in many different applications of scientific computing, such as finite element approximation to solve the Navier–Stokes equation, constrained optimization, generalized least squares problems, etc [1–3].

Since the coefficient matrix of (1) is large and sparse, iterative methods for solving (1) are effective because of storage requirements and preservation sparsity. The well-known SOR is a simple stationary iterative method, which is popular in engineering applications. One can see [12] for a comprehensive survey. Yuan [4,5] and Yuan and Iusem [6] have presented several variants of the SOR method and preconditioned conjugate gradient methods to solve general augmented systems such as (1) arising from generalized least squares problems where A can be symmetric and positive semidefinite and B can be rank deficient. In [7], Golub et al. have presented several SOR-like algorithms to solve augmented systems (1). Bai et al. [8] have presented GSOR-like algorithms to solve augmented systems (1) and obtained the optimal parameter and extended the results of [7].

Recently, Darvishi and Hessari [9] considered the following splitting:

$$\begin{pmatrix} A & B \\ -B^{\mathrm{T}} & 0 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ B^{\mathrm{T}} & 0 \end{pmatrix} - \begin{pmatrix} 0 & -B \\ 0 & Q \end{pmatrix},$$

☆ This research was supported by 973 Programs (2008CB317110), NSFC (10771030), the Scientific and Technological Key Project of the Chinese Education Ministry (107098), the Ph.D. Programs Fund of Chinese Universities (20070614001), Sichuan Province Project for Applied Basic Research (2008JY0052) and the Project for Academic Leader and Group of UESTC.

* Corresponding author. Tel.: +86 28 83201175; fax: +86 28 83200131.
E-mail addresses: wushiliang1999@126.com (S.-L. Wu), tzhuang@uestc.edu.cn, tingzhuhuang@126.com (T.-Z. Huang).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.10.006

with Q being nonsingular symmetric and applied the Symmetric SOR (SSOR) method to solve (1). Let

$$J_w = \begin{pmatrix} (1-w)^2 I - w^2(2-w)A^{-1}BQ^{-1}B^{\mathrm{T}} & -w(2-w)A^{-1}B + \dfrac{w^3(2-w)}{1-w}A^{-1}BQ^{-1}B^{\mathrm{T}}A^{-1}B \\ w(2-w)Q^{-1}B^{\mathrm{T}} & I - \dfrac{w^2(2-w)}{1-w}Q^{-1}B^{\mathrm{T}}A^{-1}B \end{pmatrix}.$$

Using the symmetric SOR method to solve augmented systems, Darvishi and Hessari [9] gave the following result.

Theorem 1 *([9]). Suppose that μ is an eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}B$. If λ satisfies*

$$(1-w)(1-\lambda)(\lambda-(w-1)^2) = w^2(w-2)^2\lambda\mu \tag{2}$$

then λ is an eigenvalue of J_w. Conversely, if λ is an eigenvalue of J_w such that $\lambda \neq (w-1)^2$, $\lambda \neq 1$ and μ satisfies (2), then μ is a nonzero eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}B$.

By observing the work of [9], it is easy to find that it is very difficult to determine its optimal iteration parameter and the corresponding optimal convergence factor from (2) in that the order of parameter w is very high. To overcome this difficulty, this paper is devoted to a splitting coefficient matrix of (1) different from [9] to reduce the order of the parameter and the optimal parameter is determined under certain conditions.

The outline of this paper is as follows. In Section 2, we present the modified symmetric successive overrelaxation (MSSOR) method to solve augmented systems. In Section 3, we study the optimal parameter for the MSSOR method defined in Section 2 and also the behavior of its spectral radius. In Section 4, we apply the MSSOR method to solve augmented systems. Finally, the paper is concluded in Section 5.

2. Modified symmetric SOR method

For the sake of simplicity, we rewrite system (1) as

$$\begin{pmatrix} A & B \\ -B^{\mathrm{T}} & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ -q \end{pmatrix}.$$

where $A \in R^{m\times m}$ is symmetric and positive definite, $B \in R^{m\times n}$. To construct the MSSOR method, we consider the follow splitting:

$$\begin{pmatrix} A & B \\ -B^{\mathrm{T}} & 0 \end{pmatrix} = D - L - U,$$

where

$$D = \begin{pmatrix} A & 0 \\ 0 & Q \end{pmatrix}, \qquad L = \begin{pmatrix} 0 & 0 \\ B^{\mathrm{T}} & \dfrac{1}{2}Q \end{pmatrix}, \qquad U = \begin{pmatrix} 0 & -B \\ 0 & \dfrac{1}{2}Q \end{pmatrix},$$

and Q is nonsingular symmetric matrix.

Let

$$z^{(k)} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix}, \qquad c = \begin{pmatrix} b \\ -q \end{pmatrix}.$$

From the symmetric SOR, we obtain the following scheme:

$$(D - wL)z^{(k+\frac{1}{2})} = [(1-w)D + wU]z^{(k)} + wc.$$

That is,

$$z^{(k+\frac{1}{2})} = \widetilde{L}_w z^{(k)} + w(D-wL)^{-1}c, \tag{3}$$

where

$$\widetilde{L}_w = (D-wL)^{-1}[(1-w)D + wU] = \begin{pmatrix} (1-w)I & -wA^{-1}B \\ \dfrac{w(1-w)}{1-\frac{w}{2}}Q^{-1}B^{\mathrm{T}} & I - \dfrac{w^2}{1-\frac{w}{2}}Q^{-1}B^{\mathrm{T}}A^{-1}B \end{pmatrix}.$$

Note that

$$D - wL = \begin{pmatrix} A & 0 \\ -wB^{\mathrm{T}} & \left(1-\dfrac{w}{2}\right)Q \end{pmatrix}.$$

Since the matrix A is symmetric and positive definite and Q is nonsingular, therefore

$$\det(D - wL) = \left(1 - \frac{w}{2}\right)^n \det(A)\det(Q) \neq 0$$

if and only if $1 - \frac{w}{2} \neq 0$, i.e., $w \neq 2$.

And so by the backward SOR, we compute $z^{(k+1)}$ from $z^{(k+\frac{1}{2})}$

$$z^{(k+1)} = \widetilde{U}_w z^{k+\frac{1}{2}} + w(D - wU)^{-1}c \tag{4}$$

where

$$\widetilde{U}_w = (D - wU)^{-1}[(1-w)D + wL] = \begin{pmatrix} (1-w)I - \dfrac{w^2}{1-\frac{w}{2}}A^{-1}BQ^{-1}B^{\mathrm{T}} & -wA^{-1}B \\ \dfrac{w}{1-\frac{w}{2}}Q^{-1}B^{\mathrm{T}} & I \end{pmatrix}.$$

From (3) and (4), we get the MSSOR iterative method as follows:

$$z^{(k+1)} = \Omega_w z^{(k)} + C,$$

with

$$\begin{aligned}\Omega_w &= \widetilde{U}_w \widetilde{L}_w \\ &= \begin{pmatrix} (1-w)^2 I - \dfrac{2w^2(1-w)}{1-\frac{w}{2}}A^{-1}BQ^{-1}B^{\mathrm{T}} & -w(2-w)A^{-1}B + \dfrac{2w^3}{1-\frac{w}{2}}A^{-1}BQ^{-1}B^{\mathrm{T}}A^{-1}B \\ \dfrac{2w(1-w)}{1-\frac{w}{2}}Q^{-1}B^{\mathrm{T}} & I - \dfrac{2w^2}{1-\frac{w}{2}}Q^{-1}B^{\mathrm{T}}A^{-1}B \end{pmatrix}\end{aligned}$$

and

$$C = w(2-w)\begin{pmatrix} A^{-1}b - \dfrac{w^2}{(1-\frac{w}{2})^2}A^{-1}BQ^{-1}B^{\mathrm{T}}A^{-1}b + \dfrac{w}{(1-\frac{w}{2})^2}A^{-1}BQ^{-1}q \\ \dfrac{w}{(1-\frac{w}{2})^2}Q^{-1}B^{\mathrm{T}}A^{-1}b - \dfrac{1}{(1-\frac{w}{2})^2}Q^{-1}q \end{pmatrix}.$$

Modified symmetric SOR(MSSOR) method: Given initial vectors $x^{(0)} \in R^n$ and $y^{(0)} \in R^m$, and a relaxation factor $w > 0$. For $k = 0, 1, 2, \ldots$ until the iteration sequence $\{(x^{(k)}, y^{(k)})^{\mathrm{T}}\}$ is convergent, compute

$$\begin{cases} y^{(k+1)} = y^{(k)} + \dfrac{4wQ^{-1}B^{\mathrm{T}}}{2-w}[(1-w)x^{(k)} - wA^{-1}By^{(k)} + wA^{-1}b] - \dfrac{4w}{2-w}Q^{-1}q \\ x^{(k+1)} = (1-w)^2 x^{(k)} - wA^{-1}B[y^{(k+1)} + (1-w)y^{(k)}] + w(2-w)A^{-1}b \end{cases}$$

where Q is an approximate (preconditioning) matrix of the Schur complement matrix $B^{\mathrm{T}}A^{-1}B$.

To study the convergence region for parameter w in the MSSOR method to solve augmented systems (2), we need the following Theorem 2.

Theorem 2. *Suppose that μ is an eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}B$. If λ satisfies*

$$(1-\lambda)(\lambda - (w-1)^2) = 4w^2\lambda\mu \tag{5}$$

then λ is an eigenvalue of Ω_w. Conversely, if λ is an eigenvalue of Ω_w such that $\lambda \neq (w-1)^2$, $\lambda \neq 1$ and μ satisfies (5), then μ is a nonzero eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}B$.

Proof. Suppose that λ and x are eigenvalue and eigenvector of Ω_w, respectively. Then, we have

$$\Omega_w x = \lambda x.$$

By calculation, we obtain that

$$\begin{aligned} &\begin{pmatrix} (1-w)^2 I & -w(1-w)A^{-1}B \\ w(1-w)Q^{-1}B^{\mathrm{T}} & \left(1-\dfrac{w}{2}\right)^2 I - w^2 Q^{-1}B^{\mathrm{T}}A^{-1}B \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \lambda \begin{pmatrix} I & wA^{-1}B \\ -wQ^{-1}B^{\mathrm{T}} & \left(1-\dfrac{w}{2}\right)^2 I - w^2 Q^{-1}B^{\mathrm{T}}A^{-1}B \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

So we get the following two equations:

$$\begin{cases} ((w-1)^2-\lambda)x_1 = w(1+\lambda-w)A^{-1}Bx_2 \\ (1-\lambda)\left(1-\dfrac{w}{2}\right)^2 x_2 - (1-\lambda)w^2Q^{-1}B^{\mathrm{T}}A^{-1}Bx_2 = w(w-1-\lambda)Q^{-1}B^{\mathrm{T}}x_1 \end{cases}$$

Hence from the first equation, we have

$$x_1 = \frac{w(1+\lambda-w)}{(w-1)^2-\lambda}A^{-1}Bx_2.$$

Taking the place of x_1 in the second equation yields

$$(1-\lambda)\left(1-\frac{w}{2}\right)^2 x_2 - (1-\lambda)w^2Q^{-1}B^{\mathrm{T}}A^{-1}Bx_2 = w(w-1-\lambda)\frac{w(1+\lambda-w)}{(w-1)^2-\lambda} \times Q^{-1}B^{\mathrm{T}}A^{-1}Bx_2.$$

By simple manipulations, it is easy to get that

$$(1-\lambda)(\lambda-(w-1)^2)x_2 = 4\lambda w^2 Q^{-1}B^{\mathrm{T}}A^{-1}Bx_2.$$

Suppose that μ is an eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}Bx_2$, then we have

$$(1-\lambda)(\lambda-(w-1)^2) = 4w^2\lambda\mu.$$

We can prove the second assertion by reversing the process. □

Remark 1. Obviously, the order of parameter w in Eq. (5) is lower than (11) of [9], which is a good signal to us for the determination of the optimum value of the parameter.

The following lemma is quoted for the latter use.

Lemma 1 *([10]). Consider the real quadratic equation $x^2-bx+d=0$, where b and d are real numbers. Both roots of the equation are less than one in modulus if and only if, $|d|<1$ and $|b|<1+d$.*

Theorem 3. *Let A and Q be symmetric positive definite, and B be of full rank. Assume that all eigenvalues μ of $Q^{-1}B^{\mathrm{T}}A^{-1}B$ are real and positive. Then we have the following cases for the relaxation parameter of the MSSOR method:*

Case 1: if $0<\mu\le\frac{1}{4}$, the MSSOR method converges for all w such that $0<w<2$.

Case 2: if $\mu=\frac{1}{2}$, the MSSOR method converges for all w such that $0<w<1$.

Case 3: if $\frac{1}{4}<\mu$ and $\mu\ne\frac{1}{2}$, the MSSOR method converges for all w such that

$$0<w<\frac{1-\sqrt{4\rho-1}}{1-2\rho} = \frac{2}{1+\sqrt{4\rho-1}} < 2,$$

where ρ is the spectral radius of $Q^{-1}B^{\mathrm{T}}A^{-1}B$.

Proof. After some simple manipulations on (5), we get

$$\lambda^2 - [1+(w-1)^2-4w^2\mu]\lambda + (w-1)^2 = 0. \tag{6}$$

By setting $b=1+(w-1)^2-4w^2\mu$ and $d=(w-1)^2$. From Lemma 1, we have

$$|(w-1)^2|<1.$$

Also, the above inequality is equal to

$$|w-1|<1.$$

Hence, it is easy to get that

$$0<w<2.$$

From $|b|<1+d$ in Lemma 1, we have

$$|1+(w-1)^2-4w^2\mu| < 1+(w-1)^2.$$

Therefore

$$-1-(w-1)^2 < 1+(w-1)^2-4w^2\mu < 1+(w-1)^2.$$

The above relation changes to the following inequalities:

$$-4w^2\mu < 0 \tag{7}$$

and

$$2 + 2(w-1)^2 - 4w^2\mu > 0. \tag{8}$$

It is easy to see that inequality (7) is true if $w \neq 0$ and $\mu > 0$.

From (8), we have

$$w^2(1-2\mu) - 2w + 2 > 0. \tag{9}$$

Case 1: If $0 < \mu < \frac{1}{4}$, from (9) it is easy to obtain that

$$\Delta = (-2)^2 - 4 \cdot 2 \cdot (1-2\mu) = 4(-1+4\mu) < 0.$$

It is not difficult to get that the MSSOR method converges for all w such that $0 < w < 2$ with $0 < \mu < \frac{1}{4}$. If $\mu = \frac{1}{4}$, we get

$$\frac{1}{2}w^2 - 2w + 2 > 0.$$

It is easy to obtain that $w \neq 2$. That is, if $\mu = \frac{1}{4}$, the MSSOR method converges for all w such that $0 < w < 2$.

Case 2: If $\mu = \frac{1}{2}$, we get

$$-2w + 2 > 0.$$

It is easy to obtain that $w < 1$. That is, if $\mu = \frac{1}{2}$, the MSSOR method converges for all w such that $0 < w < 1$.

Case 3: If $\frac{1}{4} \leq \mu < \frac{1}{2}$, by straightforwardly solving (9), we immediately obtain

$$w < \frac{1 - \sqrt{4\mu - 1}}{1 - 2\mu}.$$

Let

$$f(\mu) = \frac{1 - \sqrt{4\mu - 1}}{1 - 2\mu}.$$

It is easy to prove that $f(\mu)$ is monotone decreasing function in $[\frac{1}{4}, \frac{1}{2})$. As for any μ we have $\mu \leq \rho$, hence *Case* 3 holds.

If $\mu > \frac{1}{2}$, we get

$$w^2(2\mu - 1) + 2w - 2 < 0. \tag{10}$$

By straightforwardly solving (10), we immediately obtain

$$w < \frac{-1 + \sqrt{4\mu - 1}}{2\mu - 1}.$$

Analogously, *Case 3* holds and this completes the proof. □

3. Choice of the optimal relaxation parameter

In the section, we present the following theorems to provide the choice of the optimal relaxation parameter. For simplicity of notation, let $\rho = \rho(Q^{-1}B^{\mathrm{T}}A^{-1}B)$ and $0 < \mu_0 = \min_{\mu \neq 0} \mu$, where μ is a nonzero eigenvalue of $Q^{-1}B^{\mathrm{T}}A^{-1}B$.

Theorem 4. *If* $\mu_0 > \frac{1}{4}$, *then*

$$\rho(\Omega_w) = \begin{cases} |1-w|, & \text{if } 0 < w \leq \dfrac{2}{1+2\sqrt{\rho}}, \\ 0.5[|1+(w-1)^2 - 4w^2\rho| + w\sqrt{(1-4\rho)((w-2)^2 - 4w^2\rho)}], & \text{if } \dfrac{2}{1+2\sqrt{\rho}} \leq w \leq \dfrac{2}{1+\sqrt{4\rho - 1}}. \end{cases}$$

Moreover, the optimal parameter w_o *and* $\rho(\Omega_{w_o})$ *are given respectively by*

$$w_o = \frac{2}{1+2\sqrt{\rho}} \quad \text{and} \quad \rho(\Omega_{w_o}) = \frac{2\sqrt{\rho} - 1}{2\sqrt{\rho} + 1}.$$

Proof. From (6), it follows that

$$\lambda = 0.5[1 + (w-1)^2 - 4w^2\mu \pm w\sqrt{(1-4\mu)((w-2)^2 - 4w^2\mu)}].$$

It is not difficult to find that λ is complex if $(1-4\mu)((w-2)^2 - 4w^2\mu) \le 0$. Further, we obtain that

$$(w-2)^2 - 4w^2\mu \ge 0.$$

That is,

$$(w - 2 - 2w\sqrt{\mu})(w - 2 + 2w\sqrt{\mu}) \ge 0,$$

which is equivalent to the following form

$$((2\sqrt{\mu}-1)w + 2)((1+2\sqrt{\mu})w - 2) \le 0.$$

So λ is complex if $\frac{-2}{2\sqrt{\mu}-1} \le w \le \frac{2}{1+2\sqrt{\mu}}$ and real if $w \le \frac{-2}{2\sqrt{\mu}-1}$ or $w \ge \frac{2}{1+2\sqrt{\mu}}$. Then

$$|\lambda| = \begin{cases} |1-w|, & \text{if } 0 < w \le \dfrac{2}{1+2\sqrt{\mu}}, \\ 0.5[|1 + (w-1)^2 - 4w^2\mu| + w\sqrt{(1-4\mu)((w-2)^2 - 4w^2\mu)}], & \text{if } \dfrac{2}{1+2\sqrt{\mu}} \le w \le \dfrac{2}{1+\sqrt{4\mu-1}}. \end{cases}$$

Since $\frac{2}{1+2\sqrt{\mu}}$ is a monotonically decreasing function, it is easy to find that we get

$$\rho(\Omega_w) = \begin{cases} |1-w|, & \text{if } 0 < w \le \dfrac{2}{1+2\sqrt{\rho}}, \\ 0.5[|1 + (w-1)^2 - 4w^2\rho| + w\sqrt{(1-4\rho)((w-2)^2 - 4w^2\rho)}], & \text{if } \dfrac{2}{1+2\sqrt{\rho}} \le w \le \dfrac{2}{1+\sqrt{4\rho-1}}. \end{cases}$$

For the optimal parameter, we rewrite (5) as $g_w(\lambda) = f_w(\lambda)$, where we define

$$g_w(\lambda) = \left[\frac{\lambda - 1 + w}{w}\right]^2 \quad \text{and} \quad f_w(\lambda) = (1-4\mu)\lambda.$$

Clearly, $g_w(\lambda)$ and $f_w(\lambda)$ pass though the point $(1, 1)$ and $(0, 0)$ respectively for all w since $g_w(1) = 1$ and $f_w(0) = 0$. The straight line f_w crosses the parabolic curve g_w. Similar to the analysis in ([11], pp. 110–111) the optimal parameter w_o is the choice that guarantees that f_{w_o} is a tangent line of g_{w_o}. Using the same idea, we get

$$w_o = \frac{2}{1+2\sqrt{\rho}} > 0,$$

because $g'_w(\lambda) = 2(\lambda - 1 + w)/w^2$ and $f'_w(\lambda) = 1 - 4\mu$. We also have

$$\rho(\Omega_{w_o}) = |1 - w_o| = \frac{2\sqrt{\rho}-1}{2\sqrt{\rho}+1}. \quad \square$$

In [7], Golub et al. gave the following theorem on the choice of the optimal relaxation parameter of the SOR-like method for augmented systems.

Theorem 5 *([7]). If $\mu_0 > \frac{1}{4}$, then*

$$\rho(\mathcal{M}_w) = \begin{cases} \sqrt{1-w}, & \textit{if } 0 < w \le \dfrac{2\sqrt{\rho}-1}{\rho}, \\ 0.5[|2 - w - w^2\rho| + w\sqrt{(w\rho+1)^2 - 4\rho}], & \textit{if } \dfrac{2\sqrt{\rho}-1}{\rho} \le w \le \dfrac{4}{1+\sqrt{4\rho+1}}, \end{cases}$$

where

$$\mathcal{M}_w = (D - wL)^{-1}[(1-w)D + wU] = \begin{pmatrix} A & 0 \\ -wB^{\mathrm{T}} & Q \end{pmatrix}^{-1} \begin{pmatrix} (1-w)A & -wB \\ 0 & Q \end{pmatrix}.$$

Moreover, the optimal parameter w_b and $\rho(\mathcal{M}_{w_b})$ are given respectively by

$$w_b = \frac{2\sqrt{\rho}-1}{\rho} \le 1 \quad \textit{and} \quad \rho(\mathcal{M}_{w_b}) = \frac{|\sqrt{\rho}-1|}{\sqrt{\rho}}.$$

Here, we now give the following result, which is a criterion for choosing between the SOR-like method and the MSSOR method to solve augmented systems.

Theorem 6. *Under conditions of Theorems 4 and 5, then*

(1) $\rho(\Omega_{w_o}) > \rho(\mathcal{M}_{w_b})$, *if* $\rho > \frac{3+\sqrt{5}}{8}$;

(2) $\rho(\Omega_{w_o}) < \rho(\mathcal{M}_{w_b})$, *if* $\frac{1}{4} < \rho < \frac{3+\sqrt{5}}{8}$;

(3) $\rho(\Omega_{w_o}) = \rho(\mathcal{M}_{w_b})$, *if* $\rho = \frac{3+\sqrt{5}}{8}$.

Proof. If $\rho \geq 1$, then

$$\rho(\Omega_{w_o}) - \rho(\mathcal{M}_{w_b}) = \frac{2\sqrt{\rho}-1}{2\sqrt{\rho}+1} - \frac{\sqrt{\rho}-1}{\sqrt{\rho}} = \frac{(2\sqrt{\rho}-1)\sqrt{\rho} - (2\sqrt{\rho}+1)(\sqrt{\rho}-1)}{(2\sqrt{\rho}+1)\sqrt{\rho}} = \frac{1}{(2\sqrt{\rho}+1)\sqrt{\rho}} > 0.$$

If $\frac{1}{4} < \rho < 1$, then

$$\rho(\Omega_{w_o}) - \rho(\mathcal{M}_{w_b}) = \frac{2\sqrt{\rho}-1}{2\sqrt{\rho}+1} - \frac{1-\sqrt{\rho}}{\sqrt{\rho}} = \frac{(2\sqrt{\rho}-1)\sqrt{\rho} - (2\sqrt{\rho}+1)(1-\sqrt{\rho})}{(2\sqrt{\rho}+1)\sqrt{\rho}} = \frac{4\rho - 2\sqrt{\rho} - 1}{(2\sqrt{\rho}+1)\sqrt{\rho}}.$$

Further, we have the following conclusion:

$$\begin{cases} 4\rho - 2\sqrt{\rho} - 1 = 0, & \text{if } \rho = \dfrac{3+\sqrt{5}}{8}, \\ 4\rho - 2\sqrt{\rho} - 1 > 0, & \text{if } \dfrac{3+\sqrt{5}}{8} < \rho < 1, \\ 4\rho - 2\sqrt{\rho} - 1 < 0, & \text{if } \dfrac{1}{4} < \rho < \dfrac{3+\sqrt{5}}{8}. \end{cases}$$

In other words, we get

$$\begin{cases} \rho(\Omega_{w_o}) - \rho(\mathcal{M}_{w_b}) = 0, & \text{if } \rho = \dfrac{3+\sqrt{5}}{8}, \\ \rho(\Omega_{w_o}) - \rho(\mathcal{M}_{w_b}) > 0, & \text{if } \dfrac{3+\sqrt{5}}{8} < \rho < 1, \\ \rho(\Omega_{w_o}) - \rho(\mathcal{M}_{w_b}) < 0, & \text{if } \dfrac{1}{4} < \rho < \dfrac{3+\sqrt{5}}{8}, \end{cases}$$

which is completed. □

4. Numerical example

In this section, we give two examples to illustrate the MSSOR method to find the solution of the related augmented systems and compare the results between the MSSOR method and the SOR-like method provided [7].

In our computations of two examples, we chose the right hand side vector $(b^T, q^T)^T \in R^{m+n}$, such that the exact solution of the augmented system (1) is $(x^{(*)^T}, y^{(*)})^T = (1, 1, \ldots, 1)^T \in R^{m+n}$. All runs with respect to both the SOR-like and the MSSOR method are started from initial vector $(x^{(0)^T}, y^{(0)^T})^T = 0$, and terminated if the current iteration satisfies ERR $< 10^{-9}$, where

$$\mathrm{ERR} = \frac{\sqrt{\|x^{(k)} - x^*\|_2^2 + \|y^{(k)} - y^*\|_2^2}}{\sqrt{\|x^{(0)} - x^*\|_2^2 + \|y^{(0)} - y^*\|_2^2}},$$

with $(x^{(k)^T}, y^{(k)^T})^T$ the final approximation solution. Choosing the matrix Q, as an approximation to the matrix $B^TA^{-1}B$ in the following Examples 1 and 2, accords to the cases listed in Table 1.

Table 1
Choices of matrix Q.

Case no.	Matrix Q	Description
I	$B^{\mathrm{T}}\hat{A}^{-1}B$	$\hat{A} = \mathrm{tridiag}(A)$
II	$B^{\mathrm{T}}\hat{A}^{-1}B$	$\hat{A} = \mathrm{diag}(A)$

Table 2
The minimum positive eigenvalue μ_o of $Q^{-1}B^{\mathrm{T}}A^{-1}B$ for Example 1.

m		128	512	1152
n		64	256	576
$m+n$		128	512	1152
Case I	μ_o	0.5319	0.5088	0.5040
Case II	μ_o	0.5162	0.5044	0.5020

Table 3
Optimal parameter, spectral radius and IT for Example 1.

m			128	512	1152
n			64	256	576
$m+n$			128	512	1152
Case I	SOR	w_{opt}	0.5958	0.3657	0.2620
		$\rho(\mathcal{M}_{w_b})$	0.6358	0.7964	0.8591
		IT	62	130	200
	MSSOR	w_o	0.3081	0.1848	0.1316
		$\rho(\Omega_{w_o})$	0.6919	0.8152	0.8684
		IT	78	147	218
Case II	SOR	w_{opt}	0.4664	0.2720	0.1915
		$\rho(\mathcal{M}_{w_b})$	0.7305	0.8533	0.8992
		IT	92	191	293
	MSSOR	w_o	0.2375	0.1367	0.0960
		$\rho(\Omega_{w_o})$	0.7625	0.8633	0.9040
		IT	108	208	311

Example 1 ([7]). Consider the augmented linear system (1) in which

$$A = \begin{pmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{pmatrix} \in R^{2p^2 \times 2p^2},$$

$$B = \begin{pmatrix} I \otimes F \\ F \otimes I \end{pmatrix} \in R^{2p^2 \times p^2},$$

and

$$T = \frac{1}{h^2} \cdot \mathrm{tridiag}(-1, 2, -1) \in R^{p \times p}, \qquad F = \frac{1}{h} \cdot \mathrm{tridiag}(-1, 1, 0) \in R^{p \times p},$$

with $\otimes$ being the Kronecker product symbol and $h = \frac{1}{p+1}$ the discretization mesh-size.

For Example 1, we set $m = 2p^2$ and $n = p^2$. Hence, the total number of variables is $m + n = 3p^2$. In Table 2, we list the minimum positive eigenvalue μ_o of $Q^{-1}B^{\mathrm{T}}A^{-1}B$, for different values of m and n. In Table 3, we list w_{opt} and w_o, and the corresponding $\rho(\mathcal{M}_{w_b})$ and $\rho(\Omega_{w_o})$, the iteration number (IT), of the SOR-like and the MSSOR methods, respectively, for various problems sizes (m, n).

Example 2 ([8]). Consider the augmented linear system (1) in which

$$A = \begin{pmatrix} I \otimes T + T \otimes I & 0 \\ 0 & I \otimes T + T \otimes I \end{pmatrix} \in R^{2p^2 \times 2p^2},$$

$$B = \begin{pmatrix} I \otimes F \\ F \otimes I \end{pmatrix} \in R^{2p^2 \times p^2},$$

and

$$T = \frac{1}{h^2} \cdot \mathrm{tridiag}(-1, 2, -1) \in R^{p \times p}, \qquad F = \frac{1}{h} \cdot K \in R^{p \times p},$$

Table 4
The minimum positive eigenvalue μ_o of $Q^{-1}B^TA^{-1}B$ for Example 2.

m		128	512	1152
n		64	256	576
$m+n$		128	512	1152
Case I	μ_o	0.5305	0.5086	0.5040
Case II	μ_o	0.5160	0.5044	0.5020

Table 5
Optimal parameter, spectral radius and IT for Example 2.

m			128	512	1152
n			64	256	576
$m+n$			128	512	1152
Case I	SOR	w_{opt}	0.5692	0.3382	0.2376
		$\rho(\mathcal{M}_{w_b})$	0.6563	0.8135	0.8731
		IT	63	128	196
	MSSOR	w_o	0.2933	0.1706	0.1193
		$\rho(\Omega_{w_o})$	0.7067	0.8294	0.8807
		IT	79	146	215
Case II	SOR	w_{opt}	0.4399	0.2489	0.1722
		$\rho(\mathcal{M}_{w_b})$	0.7484	0.8666	0.9098
		IT	94	191	291
	MSSOR	w_o	0.2235	0.1250	0.0863
		$\rho(\Omega_{w_o})$	0.7765	0.8750	0.9137
		IT	111	209	310

with

$$K = (k_{ij}) \in R^{p\times p}, \qquad k_{ij} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{|i-j|^2}{8}}, \quad i, j = 1, 2, \dots, p,$$

$\otimes$ being the Kronecker product symbol and $h = \frac{1}{p+1}$ the discretization mesh-size.

For Example 2, we have $m = 2p^2$ and $n = p^2$. The total number of variables is $m + n = 3p^2$. Note that the matrix K is a highly ill-conditioned Toeplitz matrix with rapidly decaying singular values. In Table 4, we give the minimum positive eigenvalue μ_o of $Q^{-1}B^TA^{-1}B$ with the different values of m and n. In Table 5, we list w_{opt} and w_o, and the corresponding $\rho(\mathcal{M}_{w_b})$ and $\rho(\Omega_{w_o})$, the iteration number (IT), of the SOR-like and the MSSOR methods, respectively, for various problems sizes (m, n).

By observing Examples 1 and 2, it is not difficult to see that when m increases, the optimal parameter w_{opt} and w_o are deceasing, however, the corresponding optimal convergence factors $\rho(\mathcal{M}_{w_b})$ and $\rho(\Omega_{w_o})$ of the SOR-like and the MSSOR methods gradually increases. It is clear that both methods have reasonably small convergence factors, and the asymptotic convergence factor of the SOR-like method is smaller than that of the MSSOR method when the optimal parameter is employed. It is easy to know that the iteration number of the SOR-like method is smaller than that of the MSSOR method when the optimal parameter is employed. In other words, with the optimal parameter employed, the MSSOR method is less efficient than that of the SOR-like method. Therefore, we need to improve our method, such as introduce another parameter β which is similar to the idea of [8]. We will study this case in the future.

5. Conclusion

In this paper, the MSSOR method has been established to solve augmented systems, which is a simple and powerful scheme to solve (1). Convergence analysis has been given. We derived its optimal iteration parameter w. From Theorem 6 and numerical examples, it is not difficult to find that the SOR-like method is superior to the MSSOR method. So we need to further improve the MSSOR method, such as introduce another parameter β which is similar to the idea of [8]. With regard to this case, we will further study henceforth.

References

[1] H. Elman, D. Silvester, Fast nonsymmetric iteration and preconditioning for Navier–Stokes equations, SIAM J. Sci. Comput. 17 (1996) 33–46.
[2] B. Fischcher, A. Ramage, D.J. Silverster, A.J. Wathen, Minimum residual methods for augmented systems, BIT 38 (1998) 527–543.
[3] S. Wright, Stability of augmented system factorization in interior point methods, SIAM J. Matrix Anal. Appl. 18 (1997) 191–222.
[4] J.Y. Yuan, Iterative methods for generalized least squares problems, Ph.D. Thesis, IMPA, Rio de Janeiro, Brazil, 1993.
[5] J.Y. Yuan, Numerical methods for generalized least squares problems, J. Comput. Appl. Math. 66 (1996) 571–584.
[6] J.Y. Yuan, A.N. Iusem, Preconditioned conjugate gradient methods for generalized least squares problems, J. Comput. Appl. Math. 71 (1996) 287–297.
[7] G.H. Golub, X. Wu, J.Y. Yuan, SOR-like methods for augmented systems, BIT 41 (2001) 71–85.

[8] Z.Z. Bai, B.N. Parlett, Z.Q. Wang, On generalized successive overrelaxtion methods for augmented systems, Numer. Math. 102 (2005) 1–38.
[9] M.T. Darvishi, P. Hessari, Symmetric SOR method for augmented systems, Appl. Math. Comput. 183 (2006) 409–415.
[10] D.M. Young, Iteratin Solution for Large Systems, Academic Press, New York, 1971.
[11] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[12] L.A. Hageman, D.M. Young, Applied Iterative Methods, Academic Press, New York, 1981.